



Compressed Sensing with Rank Deficient Dictionaries

Hansen, Thomas Lundgaard; Johansen, Daniel Højrup; Jørgensen, Peter Bjørn; Trillingsgard, Kasper Fløe; Arildsen, Thomas; Fyhn, Karsten; Larsen, Torben

Published in:
Globecom. I E E E Conference and Exhibition

DOI (link to publication from Publisher):
[10.1109/GLOCOM.2012.6503673](https://doi.org/10.1109/GLOCOM.2012.6503673)

Publication date:
2012

Document Version
Accepted author manuscript, peer reviewed version

[Link to publication from Aalborg University](#)

Citation for published version (APA):
Hansen, T. L., Johansen, D. H., Jørgensen, P. B., Trillingsgard, K. F., Arildsen, T., Fyhn, K., & Larsen, T. (2012). Compressed Sensing with Rank Deficient Dictionaries. *Globecom. I E E E Conference and Exhibition*, 3594-3599. <https://doi.org/10.1109/GLOCOM.2012.6503673>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Compressed Sensing with Rank Deficient Dictionaries

T. L. Hansen, D. H. Johansen, P. B. Jørgensen,
K. F. Trillingsgaard, T. Arildsen, K. Fyhn and T. Larsen
Department of Electronic Systems, Aalborg University, Denmark
Email: tha@es.aau.dk

Abstract—In compressed sensing it is generally assumed that the dictionary matrix constitutes a (possibly overcomplete) basis of the signal space. In this paper we consider dictionaries that do not span the signal space, i.e. rank deficient dictionaries. We show that in this case the signal-to-noise ratio (SNR) in the compressed samples can be increased by selecting the rows of the measurement matrix from the column space of the dictionary. As an example application of compressed sensing with a rank deficient dictionary, we present a case study of compressed sensing applied to the Coarse Acquisition (C/A) step in a GPS receiver. Simulations show that for this application the proposed choice of measurement matrix yields an increase in SNR performance of up to 5 – 10 dB, compared to the conventional choice of a fully random measurement matrix. Furthermore, the compressed sensing based C/A step is compared to a conventional method for GPS C/A.

I. INTRODUCTION

The framework of compressed sensing (CS) has received much research interest in recent years, see e.g. [1] and [2]. Compressed sensing allows signals to be sampled below the Nyquist rate (subsampled) while still allowing for reconstruction. To do this, it is required that the signal has a sparse representation in some known dictionary $\Phi \in \mathbb{C}^{N \times Q}$. Then we can write the noisy signal $\mathbf{z} \in \mathbb{C}^{N \times 1}$ as

$$\mathbf{z} = \Phi \mathbf{x} + \mathbf{w}, \quad (1)$$

where $\mathbf{w} \in \mathbb{C}^{N \times 1}$ is i.i.d. zero-mean noise (typically Gaussian distributed) and $\mathbf{x} \in \mathbb{C}^{Q \times 1}$ is a sparse vector, i.e. it has very few non-zero entries, $\|\mathbf{x}\|_0 \ll Q$. A measurement matrix, $\Theta \in \mathbb{C}^{M \times N}$, which satisfies certain requirements, is chosen to obtain the sub-sampled (compressed) measurements $\mathbf{y} \in \mathbb{C}^{M \times 1}$

$$\mathbf{y} = \Theta \mathbf{z}. \quad (2)$$

As $M \ll N$ the number of samples is significantly reduced. From the sub-sampled signal \mathbf{y} , a sparse estimate of \mathbf{x} can be found by solving the optimization problem [2]:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 < \varepsilon \quad (3)$$

where we have defined the compressed dictionary $\mathbf{A} = \Theta \Phi$ and ε is a small fixed constant. Solving (3) is known as reconstruction. Several iterative greedy algorithms exist, which finds a sparse \mathbf{x} , e.g. [3], [4].

In CS it is desired that the $\mathbf{A} = \Theta \Phi$ matrix satisfies the Restricted Isometry Property (RIP), as this gives certain

guarantees for successful reconstruction. It has been shown that choosing the measurement matrix Θ as a random matrix such as Gaussian or Bernoulli matrices gives an \mathbf{A} matrix which satisfies RIP with high probability. Such measurement matrices are therefore widely deployed [5], [6]. The dictionary Φ must be chosen such that the signal of interest is sparse, when the columns of the dictionary are used as the basis of representation. Often a square non-singular matrix is used, e.g. a discrete Fourier matrix. In this case the columns of Φ constitute a basis of $\mathbb{C}^{N \times 1}$. In other cases Φ is fat ($Q > N$) and constitutes an over-complete dictionary of $\mathbb{C}^{N \times 1}$. In both cases, the dictionary has full row-rank. In this paper, we explore the application of CS to cases where the dictionary is rank deficient, i.e. it does not have full row rank and thus its columns does not span $\mathbb{C}^{N \times 1}$. Specifically, it is found that choosing the rows of the measurement matrix as a random combination of the columns in the dictionary, increases the Signal-to-Noise Ratio (SNR) in the compressed samples. It is shown by simulations that this in turn results in better reconstruction performance.

We present one example of a CS application, where a rank deficient dictionary arises; the Coarse Acquisition (C/A) step in a receiver for the Global Positioning System (GPS). CS is here applied with the hope of achieving a reduction in computational complexity. GPS is based on Code Division Multiple Access (CDMA) and in the C/A step the received signal must be correlated with a very large number of locally generated Pseudo Random Noise (PRN) sequences. We take a software defined radio approach and assume that the GPS signal is sampled above the Nyquist rate and the C/A is subsequently performed exclusively in software. The C/A step can then be posed as a sparse decomposition problem. In [7] compressed sensing was used as a preprocessor to reduce the computational complexity of sparse decompositions of audio and speech signals. This preprocessing consists of applying a measurement matrix digitally and then the sparse decomposition is obtained directly from performing the reconstruction given by (3). With this in mind, we apply CS as a preprocessor to GPS C/A. In this case a rank deficient dictionary matrix is obtained, because the columns contain the above Nyquist rate sampled signal. Another approach of applying CS to a GPS system is given in [8], where the measurement matrix is applied in analog hardware.

The paper is structured as follows. In Section II we present

a signal model for the considered dictionary matrix. From this the measurement matrix is derived in section III. Section IV describes the application of compressed sensing to GPS C/A. Numerical simulations on the GPS system are presented in Section V followed by conclusions in Section VI.

II. SIGNAL MODEL

In the following we elaborate on the signal model in (1) to clearly specify the considered scenario, in which a special structure of the measurement matrix will enhance the SNR.

The considered dictionary $\Phi \in \mathbb{C}^{N \times Q}$ can be either square ($N = Q$) or fat ($Q > N$). We start by introducing the dictionary using its SVD

$$\Phi = \mathbf{U}\Sigma\mathbf{V}^H, \quad (4)$$

where $\mathbf{U} \in \mathbb{C}^{N \times N}$ and $\mathbf{V} \in \mathbb{C}^{Q \times Q}$ are unitary matrices and $\Sigma \in \mathbb{C}^{N \times Q}$ is a matrix with the singular values in descending order on its diagonal. The column space of Φ is equal to the span of those columns in \mathbf{U} , which correspond to nonzero singular values. We now consider dictionaries with K singular values that are approximately zero. Hence the dictionary can be represented well by a low rank approximation. Had the singular values been exactly zero, the dictionary would be rank deficient. The following derivations are however also true for singular values that are approximately zero. Define the matrices $\mathbf{U}_{\text{high}} \in \mathbb{C}^{N \times (N-K)}$, $\mathbf{U}_{\text{low}} \in \mathbb{C}^{N \times K}$, $\Sigma_{\text{high}} \in \mathbb{C}^{(N-K) \times (N-K)}$ and $\Sigma_{\text{low}} \in \mathbb{C}^{K \times K}$ such that

$$\mathbf{U} = [\mathbf{U}_{\text{high}} \ \mathbf{U}_{\text{low}}], \quad \Sigma = \begin{bmatrix} \Sigma_{\text{high}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \Sigma_{\text{low}} & \mathbf{0} \end{bmatrix}, \quad (5)$$

where $\mathbf{0}$ is the zero matrix. The matrices \mathbf{U}_{high} and Σ_{high} are therefore associated with the $N - K$ singular values that are significant, and \mathbf{U}_{low} and Σ_{low} the K singular values that are approximately zero. From this we define the two subspaces S_{high} and S_{low} by the bases \mathbf{U}_{high} and \mathbf{U}_{low} . It is noted that these subspaces are orthogonal. The SVD can therefore be written as

$$\Phi = \mathbf{U}_{\text{high}}\Sigma_{\text{high}}\mathbf{V}^H + \mathbf{U}_{\text{low}}\Sigma_{\text{low}}\mathbf{V}^H. \quad (6)$$

Since singular values in Σ_{low} are approximately zero, the dictionary Φ is well approximated by only the first term in (6). In the special case where Φ is rank deficient, the second term disappears since Σ_{low} is the zero matrix.

It is now assumed that the signal of interest is perfectly represented as $\Phi\mathbf{x}$, and hence the noisy signal \mathbf{z} can be written as in (1). The vector $\mathbf{x} \in \mathbb{C}^{Q \times 1}$ is considered as a zero-mean random sparse vector, where the entries are uncorrelated and identically distributed. As the vector \mathbf{x} is sparse, each entry is distributed such that it has a value of zero with high probability. The specific distribution of the entries in \mathbf{x} is not important, as we are only interested in the covariance matrix of \mathbf{x} given by

$$\mathbf{C}_{\mathbf{xx}} = \sigma_{\mathbf{x}}^2 \mathbf{I} \in \mathbb{R}^{Q \times Q}, \quad (7)$$

where $\sigma_{\mathbf{x}}^2$ denotes the variance of each entry in \mathbf{x} and \mathbf{I} denotes the identity matrix. It is noted that although the variance is the

same for all elements in \mathbf{x} , the vector is still assumed to be sparse.

The noise vector \mathbf{w} in (1) is assumed to be i.i.d. zero-mean with covariance matrix

$$\mathbf{C}_{\mathbf{ww}} = \sigma_{\mathbf{w}}^2 \mathbf{I} \in \mathbb{R}^{N \times N}, \quad (8)$$

where $\sigma_{\mathbf{w}}^2$ is the variance of the noise samples.

III. SELECTING THE MEASUREMENT MATRIX

In the following it is shown that choosing the rows of the measurement matrix in the space spanned by the dictionary, yields a larger SNR in the compressed samples, compared to random Gaussian or Bernoulli matrices. This result is similar to the random sample kernels derived in [8] for GPS C/A using analog compressed sensing. This result is different from what is conventional in CS, where completely random matrices are typically used and the following thus presents a performance improving technique.

Consider the i 'th compressed sample which is defined according to (1) and (2) as

$$y_i = \Theta_i \Phi \mathbf{x} + \Theta_i \mathbf{w} \quad \text{for } i = 0, \dots, M-1, \quad (9)$$

where $\Theta_i \in \mathbb{C}^{1 \times N}$ denotes the i 'th row of the measurement matrix Θ . In the compressed sample y_i , the average power of the signal of interest is $\mathbb{E} [|\Theta_i \Phi \mathbf{x}|^2]$ and the noise power is $\mathbb{E} [|\Theta_i \mathbf{w}|^2]$, thus we define the SNR of the i 'th sample as

$$\begin{aligned} \text{SNR}_{y_i}(\Theta_i) &= \frac{\mathbb{E} [|\Theta_i \Phi \mathbf{x}|^2]}{\mathbb{E} [|\Theta_i \mathbf{w}|^2]} = \frac{\mathbb{E} [\Theta_i \Phi \mathbf{x} \mathbf{x}^H \Phi^H \Theta_i^H]}{\mathbb{E} [\Theta_i \mathbf{w} \mathbf{w}^H \Theta_i^H]} \\ &= \frac{\Theta_i \Phi \mathbf{C}_{\mathbf{xx}} \Phi^H \Theta_i^H}{\Theta_i \mathbf{C}_{\mathbf{ww}} \Theta_i^H} = \frac{\sigma_{\mathbf{x}}^2}{\sigma_{\mathbf{w}}^2} \frac{\Theta_i \Phi \Phi^H \Theta_i^H}{\Theta_i \Theta_i^H}. \end{aligned} \quad (10)$$

We assume that achieving a high SNR for each compressed sample yields better reconstructions. This assumption is justified by simulation in the case study of GPS C/A. Therefore we seek conditions for the row Θ_i such that SNR_{y_i} is high. Additionally, the rows of the measurement matrix Θ_i must be linearly independent such that each compressed sample contains different information.

The fraction in (10) is known as a Rayleigh quotient [9]. It attains a maximum value equal to the largest eigenvalue λ_{max} of $\Phi \Phi^H$ when Θ_i is the eigenvector corresponding to λ_{max} . The Rayleigh quotient can similarly be minimized to the smallest eigenvalue λ_{min} when Θ_i is the eigenvector corresponding to λ_{min} . In general, when Θ_i is equal to an eigenvector of $\Phi \Phi^H$, the Rayleigh quotient is equal to the corresponding eigenvalue.

Now, we write $\Phi \Phi^H$ in terms of the SVD given by (4) as

$$\Phi \Phi^H = \mathbf{U} \Sigma \mathbf{V}^H \mathbf{V} \Sigma^H \mathbf{U}^H = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^H, \quad (11)$$

where $\mathbf{\Lambda} = \Sigma \Sigma^H$. This can be seen to be the Eigenvalue Decomposition (EVD) of $\Phi \Phi^H$, with eigenvectors as columns in \mathbf{U} and eigenvalues on the diagonal of the matrix $\mathbf{\Lambda}$. The column space of $\Phi \Phi^H$ is equal to the span of those columns in \mathbf{U} , which corresponds to nonzero eigenvalues. As these are the same as the nonzero singular values of Φ (consider that

$\Lambda = \Sigma \Sigma^H$), the column space of $\Phi \Phi^H$ is thus equal to the column space of Φ . The EVD of $\Phi \Phi^H$ can also be written in terms of \mathbf{U}_{high} and \mathbf{U}_{low} similarly to (6) as

$$\Phi \Phi^H = \mathbf{U}_{\text{high}} \Lambda_{\text{high}} \mathbf{U}_{\text{high}}^H + \mathbf{U}_{\text{low}} \Lambda_{\text{low}} \mathbf{U}_{\text{low}}^H, \quad (12)$$

where $\Lambda_{\text{high}} = \Sigma_{\text{high}} \Sigma_{\text{high}}^H$ and $\Lambda_{\text{low}} = \Sigma_{\text{low}} \Sigma_{\text{low}}^H$.

Suppose now that Θ_i is chosen as an arbitrary row vector in $\mathbb{C}^{1 \times N}$. As the union of S_{low} and S_{high} spans \mathbb{C}^N any vector in $\mathbb{C}^{1 \times N}$ can uniquely be decomposed into

$$\hat{\Theta}_i = \mathbf{g}_{\text{low}} \mathbf{U}_{\text{low}}^H + \mathbf{g}_{\text{high}} \mathbf{U}_{\text{high}}^H, \quad (13)$$

where $\mathbf{g}_{\text{low}} \in \mathbb{C}^{1 \times K}$ and $\mathbf{g}_{\text{high}} \in \mathbb{C}^{1 \times (N-K)}$. Inserting $\hat{\Theta}_i$ into (10) yields

$$\begin{aligned} \text{SNR}_{y_i}(\hat{\Theta}_i) &= \frac{\sigma_{\mathbf{x}}^2}{\sigma_{\mathbf{w}}^2} \frac{\hat{\Theta}_i \Phi \Phi^H \hat{\Theta}_i^H}{\hat{\Theta}_i \hat{\Theta}_i^H} \\ &= \frac{\sigma_{\mathbf{x}}^2}{\sigma_{\mathbf{w}}^2} \frac{\mathbf{g}_{\text{low}} \Lambda_{\text{low}} \mathbf{g}_{\text{low}}^H + \mathbf{g}_{\text{high}} \Lambda_{\text{high}} \mathbf{g}_{\text{high}}^H}{\mathbf{g}_{\text{low}} \mathbf{g}_{\text{low}}^H + \mathbf{g}_{\text{high}} \mathbf{g}_{\text{high}}^H}. \end{aligned} \quad (14)$$

We reach (14) by inserting (12) and (13) and by noting that $\mathbf{U}_{\text{low}}^H \mathbf{U}_{\text{high}}$ and $\mathbf{U}_{\text{high}}^H \mathbf{U}_{\text{low}}$ are zero matrices.

We now show that $\Theta_i = \mathbf{g}_{\text{high}} \mathbf{U}_{\text{high}}^H$ yields a better SNR than using $\hat{\Theta}_i$ as normally done [5]. By noting that all eigenvalues of Λ_{high} are strictly larger than Λ_{low} , it can now be shown that

$$\begin{aligned} \text{SNR}_{y_i}(\hat{\Theta}_i) &= \frac{\sigma_{\mathbf{x}}^2}{\sigma_{\mathbf{w}}^2} \frac{\mathbf{g}_{\text{low}} \Lambda_{\text{low}} \mathbf{g}_{\text{low}}^H + \mathbf{g}_{\text{high}} \Lambda_{\text{high}} \mathbf{g}_{\text{high}}^H}{\mathbf{g}_{\text{low}} \mathbf{g}_{\text{low}}^H + \mathbf{g}_{\text{high}} \mathbf{g}_{\text{high}}^H} \\ &< \frac{\sigma_{\mathbf{x}}^2}{\sigma_{\mathbf{w}}^2} \frac{\mathbf{g}_{\text{high}} \Lambda_{\text{high}} \mathbf{g}_{\text{high}}^H}{\mathbf{g}_{\text{high}} \mathbf{g}_{\text{high}}^H} = \text{SNR}_{y_i}(\Theta_i), \end{aligned} \quad (15)$$

is satisfied if

$$\frac{\mathbf{g}_{\text{low}} \Lambda_{\text{low}} \mathbf{g}_{\text{low}}^H}{\mathbf{g}_{\text{low}} \mathbf{g}_{\text{low}}^H} < \frac{\mathbf{g}_{\text{high}} \Lambda_{\text{high}} \mathbf{g}_{\text{high}}^H}{\mathbf{g}_{\text{high}} \mathbf{g}_{\text{high}}^H}. \quad (16)$$

This inequality is satisfied since both sides of the inequality are Rayleigh quotients. Hence the left hand side has maximum equal to the largest eigenvalue in Λ_{low} and the right hand side has minimum equal to the smallest eigenvalue of Λ_{high} .

The inequality in (15) implies that using Θ_i yields a better SNR than $\hat{\Theta}_i$. Therefore Θ_i should be chosen in the subspace S_{high} . It follows that the measurement matrix can be represented as

$$\Theta = \mathbf{G} \mathbf{U}_{\text{high}}^H, \quad (17)$$

where $\mathbf{G} \in \mathbb{C}^{M \times (N-K)}$ is a matrix chosen such that $\mathbf{A} = \Theta \Phi$ satisfies RIP. It is now assumed that choosing \mathbf{G} as a random Gaussian matrix makes \mathbf{A} satisfy RIP. This is justified by the theorems from compressed sensing, which state that choosing Θ as a random Gaussian matrix makes \mathbf{A} satisfy RIP with high probability.

However, as $\mathbf{U}_{\text{high}}^H$ is often not available directly, the following two approaches can be used for selecting a suitable measurement matrix that is on the form given by (17) without using the eigenvectors of $\Phi \Phi^H$ explicitly. In the following approaches, the matrices $\mathbf{G}_1 \in \mathbb{C}^{M \times Q}$ and $\mathbf{G}_2 \in \mathbb{C}^{M \times N}$ are chosen as random Gaussian matrices.

- 1) It is noted that the space spanned by S_{high} is approximated by the column space of Φ . Hence an appropriate choice of the measurement matrix is

$$\Theta = \mathbf{G}_1 \Phi^H \quad (18)$$

$$= \underbrace{\mathbf{G}_1 \mathbf{V} \Sigma_{\text{low}}}_{\mathbf{G}_{\text{low}}} \mathbf{U}_{\text{low}}^H + \underbrace{\mathbf{G}_1 \mathbf{V} \Sigma_{\text{high}}}_{\mathbf{G}_{\text{high}}} \mathbf{U}_{\text{high}}^H. \quad (19)$$

This matrix is approximately on the form given by (17) since the entries \mathbf{G}_{low} are low compared to \mathbf{G}_{high} . This choice of measurement matrix is used in the case study in Section IV.

- 2) By choosing the measurement matrix as

$$\Theta = \mathbf{G}_2 \mathbf{U}_{\text{high}} \mathbf{U}_{\text{high}}^H, \quad (20)$$

the compressed dictionary is given by

$$\begin{aligned} \mathbf{A} &= \Theta \Phi = \mathbf{G}_2 \mathbf{U}_{\text{high}} \mathbf{U}_{\text{high}}^H (\mathbf{U}_{\text{high}} \Sigma_{\text{high}} + \mathbf{U}_{\text{low}} \Sigma_{\text{low}}) \mathbf{V}^H \\ &= \mathbf{G}_2 \mathbf{U}_{\text{high}} \Sigma_{\text{high}} \mathbf{V}^H \approx \mathbf{G}_2 \Phi, \end{aligned} \quad (21)$$

where the last approximation follows since the dictionary is low rank approximated well by only the significant eigenvalues in Σ_{high} . The matrix \mathbf{A} now has a structure similar to what is normal in CS. It is noted that $\mathbf{U}_{\text{high}} \mathbf{U}_{\text{high}}^H$ can be formed by the matrix product $\mathbf{U}_b \mathbf{U}_b^H$ where \mathbf{U}_b is any orthonormal basis of S_{high} .

The proposed measurement matrix can be interpreted as follows; when calculating the compressed samples \mathbf{y} , the Nyquist samples \mathbf{z} are orthogonally projected onto the subspace S_{high} whereafter the compressed measurements are generated by multiplication with a random fat matrix \mathbf{G}_2 . In the projection, all signal energy in the subspace S_{low} is removed. Since most energy of the signal of interest is in the subspace S_{high} only a small amount of signal energy is removed. On the other hand, the noise energy is equally spread onto all eigenvectors \mathbf{U} , hence a large amount of noise energy is removed.

IV. APPLICATION TO COARSE ACQUISITION STEP OF A GPS RECEIVER

In this section we investigate the application of the CS framework to the C/A step of a software defined GPS receiver. This is an example of a system in which the dictionary becomes rank deficient, because the signal is oversampled. See [10] for an implementation of such a receiver. The ultimate goal of applying CS in a GPS receiver, is to reduce the computational requirements of the C/A step. In this paper we are however not investigating the computational complexity in detail, as we are more concerned with whether it is possible to apply CS to this application and what are the effects of the choice of measurement matrix.

A. GPS Signal Model

Each satellite transmits two signals designated as L1 and L2 with carrier frequencies of 1572.42 MHz and 1227.6 MHz respectively. The L2 signal is used for military purposes and is therefore not considered in the following.

For the L1 signal each satellite is assigned a unique Gold code [11], which in GPS is known as C/A codes. These C/A codes exhibit low correlation between each other and between time-shifts of the same code. Each satellite transmits a low-rate (50 Hz) data signal, which is modulated by the high-rate (1.023 MHz) C/A code. As the C/A codes are of length 1023, they are repeated every 1 ms. There are 32 different C/A codes in use but at any given time only a subset of these are within the field of view of the GPS receiver [12]. Due to a high velocity difference between the satellites and the receiver, a significant Doppler shift is introduced. The satellites and receiver are not synchronized in time, hence the phase of each transmitted C/A code (code phase) is unknown. The purpose of the C/A step is to identify which satellites are within the field of view and to determine the Doppler shift and code phase of these.

Navigation data and the C/A codes are modulated onto the in-phase component of the L1 frequency, while navigation data and the so-called P-code are modulated onto the quadrature component of the same carrier frequency. The P-code is a pseudo random noise sequence and its transmit power is 3 dB lower than that of the C/A code. For the C/A step of a GPS receiver only the C/A code is of any interest and the P-code is therefore considered as co-channel noise. As the power of the P-code is typically lower than the noise floor of the receiver, it is ignored. When needed, the P-code can be extracted from the noise due to the processing gain of CDMA. The navigation data modulated onto the in-phase component is neither included in the model. It is thus assumed that no navigation data shifts occur during the C/A step. This is fair as the duration of one navigation data bit is much longer than the time-span over which the C/A step is performed. In a real receiver there exist techniques to handle the situation where a data shift occur in the signal-duration where the C/A step is performed [10].

The ideally received baseband signal from the s 'th satellite is thus modelled to solely consist of the C/A code spreading waveform, denoted as $g^{(s)}(t)$. It is periodic with 1 ms and can be written as

$$g^{(s)}(t) = \sum_{k=-\infty}^{\infty} \hat{g}^{(s)}(t - kT_{\text{chip}}N_1), \quad (22)$$

where $\hat{g}^{(s)}$ is the contribution from one period, $N_1 = 1023$ is the length of the C/A code and $T_{\text{chip}} = \frac{1}{1.023 \text{ MHz}}$ is the duration of one chip. The contribution from one period is given by

$$\hat{g}^{(s)}(t) = \sum_{n=0}^{N_1-1} c^{(s)}[n] \cdot v(t - nT_{\text{chip}}), \quad (23)$$

where $c^{(s)}[n] \in \{-1, 1\}$ is the bipolar C/A code of length N_1 , and $v(t)$ is the impulse response of the channel including transmit and receive filters. Assuming a signal linear receiver, its frequency response is given by

$$V(f) = R_t(f) \cdot R_x(f) \cdot H(f), \quad (24)$$

where the transmit filter $R_t(f)$ is modeled as an ideal sharp cut-off low-pass filter with a one-sided bandwidth of at least 10.23 MHz [12], and the channel frequency response $H(f)$ is constant since multipath effects are not considered. The receive filter $R_x(f)$ is dominating and is modeled as an ideal low-pass filter with a one-sided bandwidth equal to half the sampling frequency f_s . It is also acting as anti-aliasing filter.

Under the effects of channel attenuation, Doppler shift, time delay, Additive White Gaussian noise (AWGN) and after downconversion to baseband, the summed signal from all satellites in complex baseband representation is given by

$$z(t) = \sum_{s=0}^{S-1} \alpha^{(s)} g^{(s)}(t - t_{\text{cd}}^{(s)}) \exp \left[j \left(\omega_d^{(s)} t + \theta^{(s)} \right) \right] + w(t), \quad (25)$$

where $\alpha^{(s)}$ is the channel attenuation from the s 'th satellite to the receiver, $t_{\text{cd}}^{(s)}$ is the time delay or code phase of the C/A code for the s 'th satellite, $\omega_d^{(s)}$ is the Doppler frequency for the s 'th satellite, $\theta^{(s)}$ is the phase of the carrier at the receiver and $w(t)$ is complex-valued additive white Gaussian noise.

B. Coarse Acquisition as a Sparse Decomposition Problem

The problem of C/A is to determine the Doppler shift $\omega_d^{(s)}$ and the code phase $t_{\text{cd}}^{(s)}$ in (25) for each satellite. The signal model does not have an obvious sparse structure that can be exploited for formulation of the sparse decomposition problem since the parameters $t_{\text{cd}}^{(s)}$ and $\omega_d^{(s)}$ are considered unknown, continuous-valued constants. To identify a sparse structure these values are quantized similarly to the procedure in [8]. Denote the number of different Doppler shifts as D and the number of different equally spaced code phases as L . In worst case the Doppler shift can deviate up to ± 10 kHz, and it is in most cases sufficient to know the Doppler frequency in steps of 500 Hz [10]. In this case $D = 2 \frac{10000}{500} + 1 = 41$. For a code phase resolution of 1 chip, $L = 1023$ which is the length of the C/A code. The signal model is approximated as follows

$$z(t) \approx \sum_{s=0}^{S-1} \sum_{d=0}^{D-1} \sum_{\ell=0}^{L-1} g^{(s)}(t - \ell T_c) \cdot \exp \left[j \left(\mathcal{D}_d \cdot \Delta\omega \cdot t + \theta^{(s)} \right) \right] \cdot \alpha^{(s,d,\ell)} + w(t), \quad (26)$$

where $\mathcal{D} = \{-20, -19, \dots, 19, 20\}$ is the set defining possible Doppler shifts along with the step size $\Delta\omega = 2\pi 500 \frac{\text{rad}}{\text{s}}$, $T_c = \frac{1 \text{ ms}}{L}$ is the step size for the code phase and $\alpha^{(s,d,\ell)}$ is the amplitude associated with the s 'th satellite, the d 'th Doppler shift and the ℓ 'th code phase. Note that $\alpha^{(s,d,\ell)}$ only has one non-zero element for each s , because the amplitude of each satellite only pertains to one Doppler shift and one code phase. This is what makes the signal have a sparse representation. Rewriting (26) as

$$z(t) \approx \sum_{s=0}^{S-1} \sum_{d=0}^{D-1} \sum_{\ell=0}^{L-1} \phi^{(s,d,\ell)}(t) \cdot x^{(s,d,\ell)} + w(t), \quad (27)$$

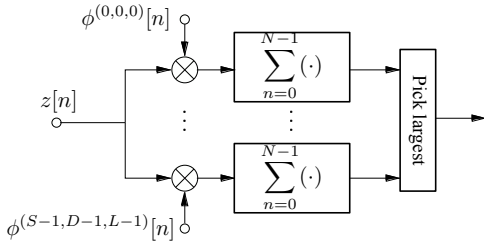


Fig. 1. Conceptual illustration of coarse acquisition. The received signal $z[n]$ is correlated with the C/A codes under all combinations of code phases and Doppler-shifts. The number of correlators is SDL , where S is the number of C/A codes, D is the number of Doppler-shifts and L is the number of code phases considered. The largest correlation values signify the present satellites, their Doppler-shift and code phase values.

where

$$\phi^{(s,d,\ell)}(t) = g^{(s)}(t - \ell T_c) \cdot \exp[j\mathcal{D}_d \cdot \Delta\omega \cdot (t - \ell T_c)], \quad (28)$$

$$x^{(s,d,\ell)} = \alpha^{(s,d,\ell)} \cdot \exp[j\mathcal{D}_d \cdot \Delta\omega \cdot \ell T_c] \cdot \exp[j\theta^{(s)}], \quad (29)$$

shows that the approximation of $z(t)$ is sparse with respect to the basis functions $\phi^{(s,d,\ell)}(t)$, since $\alpha^{(s,d,\ell)}$ is only nonzero for a few triplets (s, d, ℓ) .¹ The GPS C/A problem can be solved by finding the triplets for which $\alpha^{(s,d,\ell)}$ is nonzero.

Since it is desired to solve the problem in the digital domain, we sample the signal $z(t)$ with a sample frequency $f_s = \frac{1}{T_s}$ and thereby define the following vectors:

$$\mathbf{z} = [z(t_0), \dots, z(t_{N-1})]^T \in \mathbb{C}^{N \times 1} \quad (30)$$

$$\boldsymbol{\Phi}^{(s,d,\ell)} = [\phi^{(s,d,\ell)}(t_0), \dots, \phi^{(s,d,\ell)}(t_{N-1})]^T \in \mathbb{C}^{N \times 1} \quad (31)$$

$$\mathbf{w} = [w(t_0), \dots, w(t_{N-1})]^T \in \mathbb{C}^{N \times 1}, \quad (32)$$

where N is the number of obtained samples and $t_n = n \cdot T_s$. Denote the dictionary matrix $\boldsymbol{\Phi} \in \mathbb{C}^{N \times SDL}$ that has columns given by $\boldsymbol{\Phi}^{(s,d,\ell)}$ for all combinations of s, d and ℓ

$$\boldsymbol{\Phi} = [\boldsymbol{\Phi}^{(0,0,0)}, \boldsymbol{\Phi}^{(0,0,1)}, \dots, \boldsymbol{\Phi}^{(S-1,D-1,L-1)}]. \quad (33)$$

The dictionary $\boldsymbol{\Phi}$ is fat since $SDL \gg N$. The sparse signal model can now be expressed in matrix form:

$$\mathbf{z} \approx \boldsymbol{\Phi} \mathbf{x} + \mathbf{w}, \quad (34)$$

with sparse $\mathbf{x} \in \mathbb{C}^{SDL \times 1}$ given by

$$\mathbf{x} = [x^{(0,0,0)}, x^{(0,0,1)}, \dots, x^{(S-1,D-1,L-1)}]^T. \quad (35)$$

The naive way to find \mathbf{x} is by correlating the received signal with the SDL combinations of C/A codes, Doppler frequencies and code phases as depicted in Fig. 1. This method is computationally heavy due to the large number of correlators ($SDL = 1342176$). The computations can be performed more efficiently using a Fast Fourier Transform (FFT) based approach known as Parallel Code Phase Search (PCPS) [10].

¹Implementation detail: The inclusion of $-j\mathcal{D}_d \cdot \Delta\omega \cdot \ell T_c$ into (28) allows us to write a dictionary as a concatenation of partial circulant matrices (we define a circulant matrix as in [13]). A circulant matrix is diagonalizable by the Fourier matrix, which facilitates faster matrix-vector multiplication. Solving the problem in (3) iteratively can thus be implemented more efficiently.

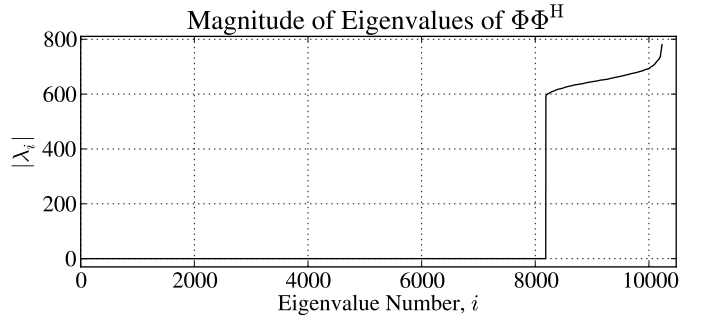


Fig. 2. Magnitude of eigenvalues (sorted) of $\boldsymbol{\Phi}\boldsymbol{\Phi}^H$, for a dictionary $\boldsymbol{\Phi}$ generated according to (33) with $S = 32$, $D = 41$, $L = 1023$, $N = 10230$ and sampling frequency $f_s = 5.115$ MHz.

V. SIMULATIONS & RESULTS

In order to assess the performance of the proposed measurement matrix and the compressed sensing based C/A scheme we carry out a numerical simulation². First we show that the dictionary is indeed rank deficient. To do this, the signal is sampled with a sample frequency 5 times the chipping rate, $f_s = 5 \cdot 1.023$ MHz. We sample two periods of the C/A code which corresponds to $N = 10230$ samples. From this the dictionary is constructed and the eigenvalues of $\boldsymbol{\Phi}\boldsymbol{\Phi}^H$ are calculated. The magnitudes of these eigenvalues are shown in Fig. 2. It is clearly seen that only a few of the eigenvalues of $\boldsymbol{\Phi}\boldsymbol{\Phi}^H$ are significant. From (11) follows that the singular values of $\boldsymbol{\Phi}$ is the square root of the eigenvalues of $\boldsymbol{\Phi}\boldsymbol{\Phi}^H$. Therefore, only a few of these singular values are significant and $\boldsymbol{\Phi}$ is thus rank deficient.

To solve the optimization problem in (3) efficiently, we use the state-of-the-art reconstruction algorithm CoSaMP [4] elaborated with the concept of model-based CS [14]. The model-based approach is basically restricting the CoSaMP algorithm to not produce more than one nonzero element of \mathbf{x} for each C/A-code. To reduce the computational requirements for the simulations all random matrices are implemented as random circulant matrices. That is, the first column is generated as random Gaussian and the remaining columns are found as circulant shifts of this. In [15] simulations show that in many cases this does not affect the reconstruction performance.

Throughout all numerical experiments the performance of the system is evaluated through its ability to correctly identify present satellites and their corresponding Doppler shift and code phase. We define the success rate performance metric as the ratio of correctly identified satellites to the total number of satellites present. A satellite is correctly identified if its estimated Doppler frequency is $0.6 \cdot \Delta\omega = 300$ Hz from the reference frequency and its estimated code phase is within $0.6 \cdot T_{\text{chip}} \approx 587$ ns from the reference code phase. The factor of 0.6 is somewhat arbitrarily chosen, such that two adjacent search steps are considered successful, when the true parameter falls in between these two.

²To comply with the reproducible research paradigm, the (Python based) source code used for the simulations can be found online at <http://sparsesampling.com/globecom2012>

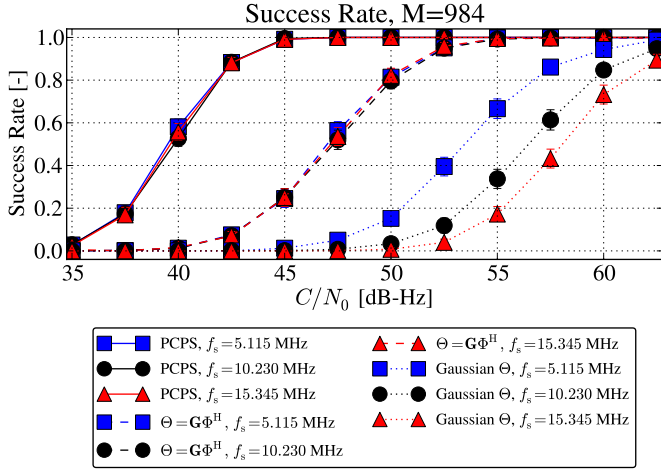


Fig. 3. Success rate versus $\frac{C}{N_0}$ for different Nyquist sample frequencies. For the compressed sensing based approach, 984 compressed samples are used. Parallel Code Phase Search (PCPS) is utilizing all Nyquist samples from a 2 ms signal. Approximated 99% confidence intervals are shown.

The used measure for channel noise is the carrier-to-noise-density ratio $\frac{C}{N_0}$ [Hz], where C is the average power of the passband signal and N_0 is the noise spectral density. Typically $\frac{C}{N_0}$ ranges from 35 to 55 dBHz [16]. The signal from one satellite is generated in complex baseband with an average sample power P_s . The noise is a zero-mean discrete white process where the amplitudes have a complex Gaussian distribution. It can be shown that the variance σ_W^2 of the complex noise is given by

$$\frac{P_s}{\sigma_W^2} = \frac{C}{N_0} \frac{1}{f_s}. \quad (36)$$

For each simulation the signals with equal power is generated for 8 different satellites randomly chosen from the set $\{0, \dots, 31\}$. When running the simulations the reconstruction algorithm is given the number of satellites present in the signal. Each data point is averaged over 100 simulations, each with a new randomly generated signal and measurement matrix Θ .

To illustrate the effect of choosing a measurement matrix according to the result in Section III, two choices of the measurement matrix are used; $\Theta = \mathbf{G}\Phi^H$ and Θ equal to a random Gaussian matrix of dimensions $M \times N$ with independent entries. The result of this simulation is shown for different sampling frequencies in Fig. 3. First of all it is noted that the conventional method of Parallel Code Phase Search (PCPS) performs equally well for all simulated sampling frequencies. The same is the case for $\Theta = \mathbf{G}\Phi^H$. However when using Θ equal to a Gaussian matrix, the ability to identify the correct support of the signal decreases with the sampling frequency. The proposed measurement matrix thus significantly improves performance and this is expected to generalize for all cases where rank deficient dictionaries are used. It is noted that the proposed compressed sensing based approach to GPS C/A does not perform as well as PCPS. The proposed approach might however still find applications in receivers which are known to operate in the high SNR

region, if the computational requirements can be reduced. In this paper we have not explored the very important aspect of the computational requirements for the proposed method compared to conventional methods such as PCPS.

VI. CONCLUSION

In this paper we have considered rank deficient dictionaries in the context of compressed sensing. We have shown that to increase the SNR in the compressed samples, the rows of the measurement matrix have to be chosen within the subspace spanned by the dictionary. Through a case study of compressed sensing applied to the C/A step in a GPS receiver, the proposed measurement matrix is shown to increase performance compared to the usual choice of a random measurement matrix. This is expected to generalize for all applications of compressed sensing with rank deficient dictionaries.

The case study demonstrated that compressed sensing can be applied to the C/A step of a GPS receiver. It is subject of further research to investigate if a reduction of computational requirements can also be achieved.

REFERENCES

- [1] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.
- [2] D. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [3] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressed sensing," *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, pp. 265–274, 2009.
- [4] D. Needell and J. Tropp, "Cosamp: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, 2009.
- [5] E. Candes and M. Wakin, "An introduction to compressive sampling," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 21–30, Mar. 2008.
- [6] E. Candes and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.
- [7] M. Christensen, J. Østergaard, and S. Jensen, "On compressed sensing and its application to speech and audio signals," in *Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers*, Nov. 2009, pp. 356–360.
- [8] X. Li, Y. Eldar, and A. Scaglione, "Low complexity acquisition of GPS signals," in *Signal Processing Advances in Wireless Communications (SPAWC), 2011 IEEE 12th International Workshop on*, Jun. 2011, pp. 56–60.
- [9] L. N. Trefethen and D. Bau, *Numerical Linear Algebra*. SIAM: Society for Industrial and Applied Mathematics, 1997.
- [10] K. Borre, D. M. Akos, N. Bertelsen, P. Rinder, and S. H. Jensen, *A Software-Defined GPS and Galileo Receivers - A Single-Frequency Approach*. Birkhauser, 2007.
- [11] R. Gold, "Optimal binary sequences for spread spectrum multiplexing," *IEEE Transactions on Information Theory*, vol. 13, no. 4, pp. 619–621, Oct. 1967.
- [12] *Interface Specification IS-GPS-200*, Global Positioning System Wing (GPSW) Systems Engineering & Integration Std., Rev. E, Jun. 2010.
- [13] P. J. Davis, *Circulant Matrices*. Chelsea Pub Co, 1994.
- [14] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, "Model-based compressive sensing," *IEEE Transactions in Information Theory*, vol. 56, pp. 1982–2001, Apr. 2010.
- [15] W. Yin, S. Morgan, J. Yang, and Y. Zhang, "Practical compressive sensing with Toeplitz and circulant matrices," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 7744, Jul. 2010.
- [16] M. S. Braasch and A. J. V. Dierendonck, "GPS receiver architectures and measurements," in *Proceedings of the IEEE*, vol. 87, no. 1, Jan. 1999, pp. 48–64.